

VPI Player : An Approach for Virtual Percussion Instrument using Kinect

Priti Sanap, Shabda Dongaonkar

G H Raison College of Engineering & Management, Wagholi, Pune, M.H. India

Corresponding Email: priti.ramesh.sanap@gmail.com

Abstract : *Today the HCI application become more popular due to increasing use of gesture recognition which easily removes mechanical devices for operating system and helpful for human being to operate system naturally. A VPI player is way to check the skills of all level of instrument player including professional, non-professional, new learner. VPI player captures the hand movements of person and by using information of captured gesture it produce sound of selected instrument. In this paper, the proposed system is real time music system emphasis on percussion instrument only which introduces kinect as image acquisition device which eliminates lighting condition, cluttered background. Use of cosine algorithm to calculate difference between the hands coordinates and instrument coordinates. VPI player is most promising way to produce sound through hand gesture with increased accuracy and efficiency.*

Index Terms: Kinect sensor, Hand Gesture, HCI, VPI, Cosine Algorithm.

I. Introduction

As per research report over \$15 billion Cr is estimated investment from 2013-2018 for gesture recognition techniques and electronic applications contribute to have 99% global gesture recognition market. Basically hand gesture has been used as sign language to recognize expressions and alphabets. Day by day the people are more attracted towards computer-human interaction [iv][v] applications. This need turns to have something should be invented in music world that fulfil the HCI requirement and prove itself in music industry to have an outstanding approach for instrument player. Different technologies used to gather the gesture by data glove, web camera and also coloured marker [vii] but they suffered from problem like light emitting condition, complex background, and time delay in recognition, efficiency, and robustness[viii]. Although kinect sensor used from 2010 having great achievement in gaming application [vi]; from 2012 kinect use itself for windows 8[xii]. VPI player uses kinect sensor for capturing the dynamic [x] (sequence of poses) gesture images. The percussion family is classified into tuned and Pitched percussion and unpitched percussion. [xii] One or more pitches of sound are provided by Pitched percussion; whereas indefinite pitch is produced by untuned percussion instrument. Pitch is assets of sounds that grant us ordering on a frequency-related scale, or to evaluate sounds as "higher" and "lower" in the sense associated with musical melodies. Examples are Drum, Bass Drum, Castanets triangles, and cymbals.

The Microsoft provides [xii] kinect sdk tool as open source to people to make use of available function to shape up the

application as per user need .this thesis mark on kinect capturing images and skeleton tracking .The whole skeleton is track by kinect itself ,we use only the coordinates of hands have great impact on obtaining pixels to draw players skeleton which becomes the base for gesture recognition module.

II. Literature Survey

In [iii], YU Bo ,CHEN YongQiang, HUANG Ying-Shu,XIA Chenjie ,emphasizes on the static hand gesture .here for hand gesture recognition the finger angle characteristics like fingertips and angle sizes are used. Images are captured from the webcam instead of data glove. It helps to differentiate between the static and dynamic characteristics.

In[ii], Ing-Jr Ding ,Che-Wei Chang and Chang-Jyun He they use dynamic time wrapping (DTW),Hidden Markov Model(HMM) ,principal component analysis (PCA) for recognizing the gesture captured through kinect sensor. The recognized gesture instruction is used to control the humanoid robot. In this paper they represents human action is learn easily by humanoid robot .

In [i], Prateem Chakraborty, Prashant Sarawgi,Ankit Mehrotra, Gaurav Agrawal , Ratika Pradhan[2008].they explained the different methods of hand gesture recognition like subtraction ,Principal Component Analysis, Rotation Invariant and Gradient. Here different images database are taken for four different gesture and each image before processed converted into the 8 bit gray scale images and filtering is provided to minimize the noise present in image.

In [ix], Sunny Amatya, Somrak Petchartee Kinect comes with a SDK that has some predefined library, coding available which can be used for working with different languages like XAML,Visual Basic, C++, and C.Skeleton tracking code samples are used for tracking features to sketch each skeleton data in right position. After closed hand detection extraction of the depth data of the skeleton point is done. The bone orientation library provides the position of the joints. After open hand is detected, the joint orientation of the wrist is done with respect to the parent bone to and the degree of rotation in x, y and z axis. Lasso mode of the hand detection is used to detect the pinch mode. There is pinch forward and open pinch mode in the section. This paper is proposing a real time, gesture based robotic arm manipulation using kinect sensor.This method uses kinect depth data, skeleton data and joint orientation data for end user movement including roll, pitch and pinch.

In [xi], Zhou Ren,Junsong Yuan,Jingjing Meng,Zhengyou Zhang they used kinect sensor for gathering the images. As traditional

data glove and vision based hand gesture recognition methods are far away to satisfy the need of real life applications. They uses FEDM Finger Earth Mover Distance it is robust to orientation, scaling. Optical sensors are affects the performance of hand gesture recognition whereas kinect sensor is reliable and not affected by lighting condition as well as cluttered background.

III.VPI Player System

Now people can enjoy music by traditional instrument player which is bulky and heavy to carry as well as require extra cost for moving instrument from one place to another. This thesis proposed work related to VPI player which is used to produce sound without using the heavy and bulky instruments, none of any external devices are used to interaction with this system. A kinect is useful to put input and play music as per the selected instrument. The system model of VPI player is given in fig.1.the VPI player model is divided into four module in sequence hand gesture Grabbing, skeleton tracking, gesture recognition and coordinate mapping , gesture post processing.

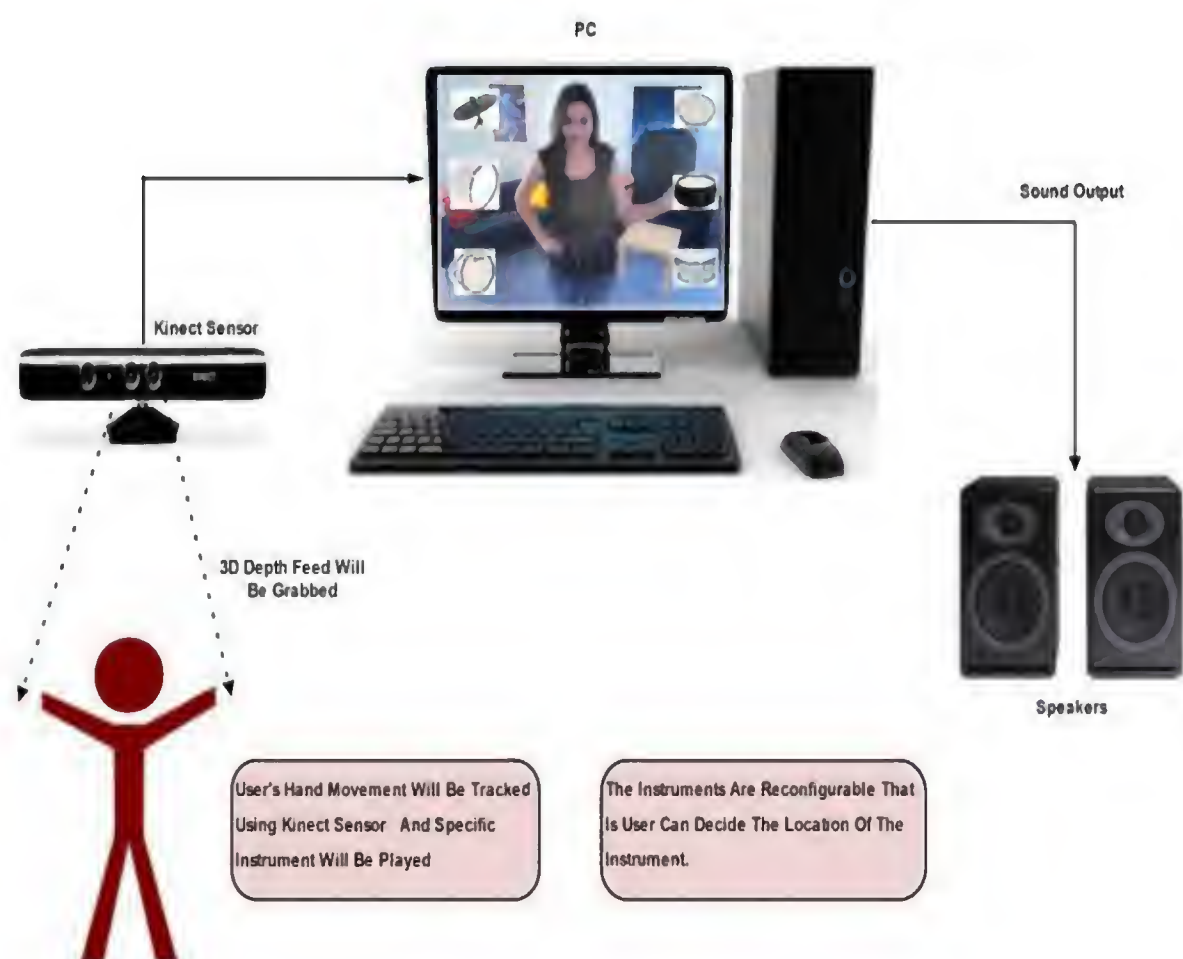


Fig.1 VPI Payer Model

A. System model

The kinect camera captures the hand gesture which is input for the VPI player. Different percussion instrument is place on screen and the particular instrument is selected depending upon the gesture position. Based on the hand movement the selected instrument will deliver the sound files for respective percussion instrument is stored in database. Basically we are not implemented any feature extraction algorithm to separate the hand gesture from outside background. Here Microsoft kinect sensor is capable of performing this task internally. The system model of the VPI player shown as in fig.1.The skeleton tracking is perform after the capturing the image.The gestures are compared with the data available with us in database. Ultimately the gesture recognizes by kinect and music is played according to the result of the cosine result of selected instrument. The cosine algorithm is useful to calculate distance between the obtain gesture coordinates to the kinect.

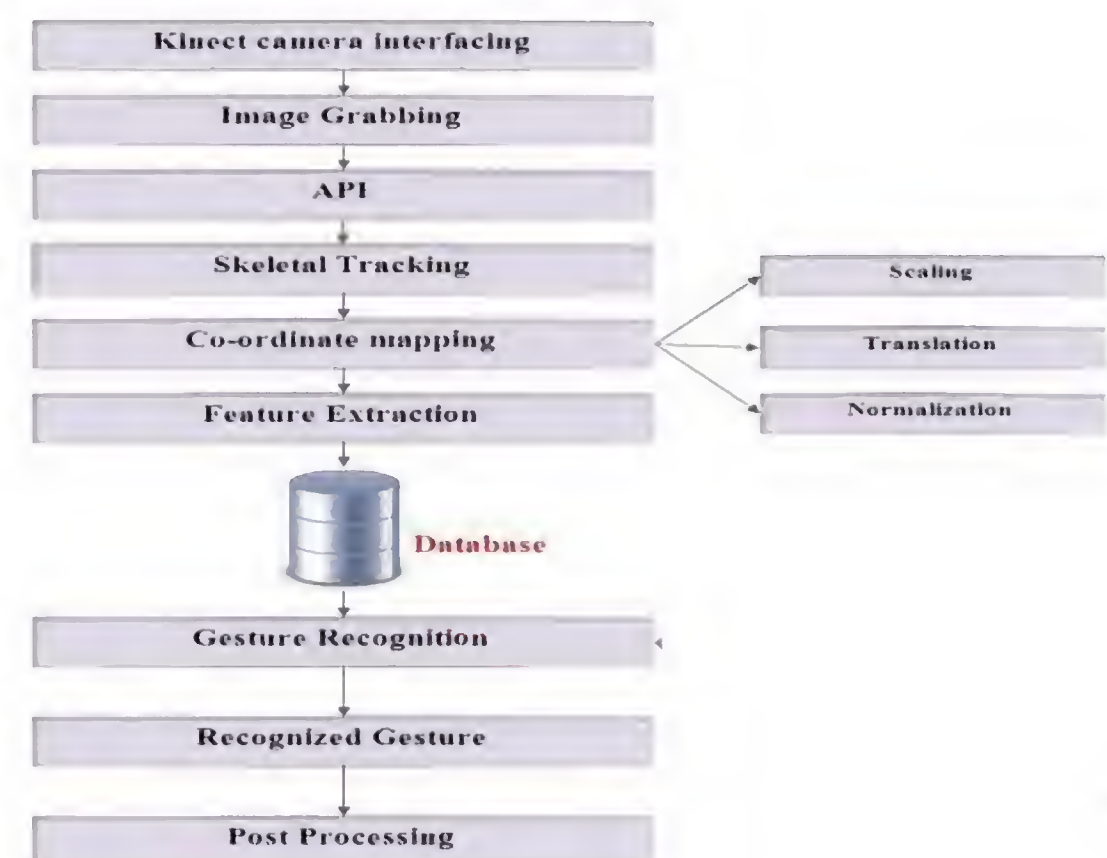


Fig.2 VPI Player System Flow

B. Kinect camera interfacing and Image Grabbing

Basically this module covers the connection of kinect to the computer system, as the xbox kinect reused here for VPI player it need the adapter to have connection between computer and kinect .start and stop buttons are provided to capture the movement of skeleton. When kinect is on, the images of hand gestures is capture using Graphics 2D function and store the frame as buffer image. It is feed to next module for skeleton tracking. Due to the use of z axis it is possible to capture the depth of image as well.

C. Skeleton tracking

It is possible to identify people and pursue their actions. In the line of view of sensor the users are identified by kinect with the help of infrared (IR) camera, For this thesis kinect track the skeleton of person who is standing in front of kinect for identifying hand movement to play the percussion instrument. The VPI player can pinpoint the joints of the tracked users in space also helpful to track their movements respect with time. Kinect identified standing or sitting, hand movement of user. Sometimes sideways poses creates some challenges with respect to the part of the user which not visible to the sensor. To recognize the user should in front of kinect sensor and make sure that sensor can get the body; here not any calibration or pose is needed for tracking. By kinect it is possible to capture 6 skeletons at a time but this system captures only one which obtained at first. The skeleton tracked by kinect is fit into specified window. If person is not stand properly in front of kinect then system will take time for capturing skeleton.

D. Coordinate mapping

The grabbed image is feed for skeleton tracking and scaling, normalization, translation operation performed by this module to have the image in predefined area. So that the image could not be scattered over the Screen, it helpful to avoid distorted pixel image. Basically we get hands coordinates as well as instrument coordinates. Now main purpose is which instrument should played it is decides by the difference between hand coordinates and instrument coordinates calculated using cosine algorithm. The result is compare with predefined threshold value .the comparison result suggests playing particular instrument. If the hand strikes faster than difference between closest coordinate and last coordinate is calculated to control pitch of sound similarly for slower stroke here difference is calculated between initial coordinate and last coordinate. Depending on selected instrument the sound file is played which store in database. Likewise it is possible to reconfigure other percussion instrument as per users need as well as customization of sound for greater background music can be achievable with VPI player.

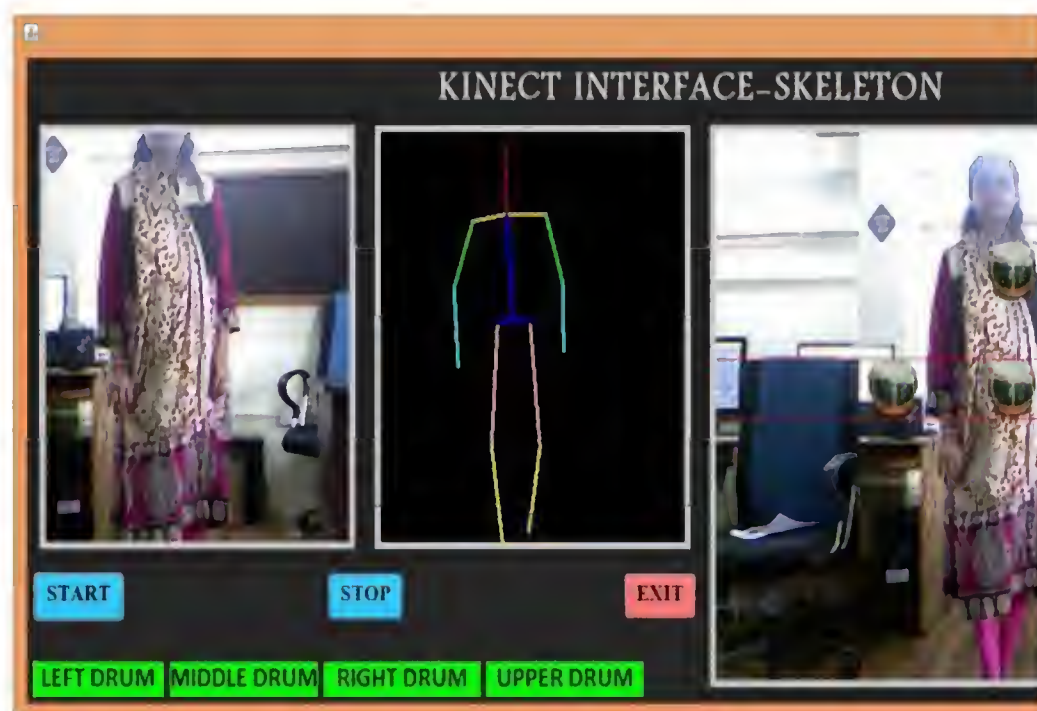


Fig3.VPI Player with overlapping window

IV Performance Evaluation

A. Precision and Recall

This system require basic setup including kinect sensor ,the adapter connecting kinect to computer ,Intel Core i5540M processor for windows 8 with 8GB RAM. The performance is calculated in terms of precision and recall factor .Where precision is ratio of recognized beats with capture gesture. Recall is the ratio of accurately performed beats per obtained gestures.

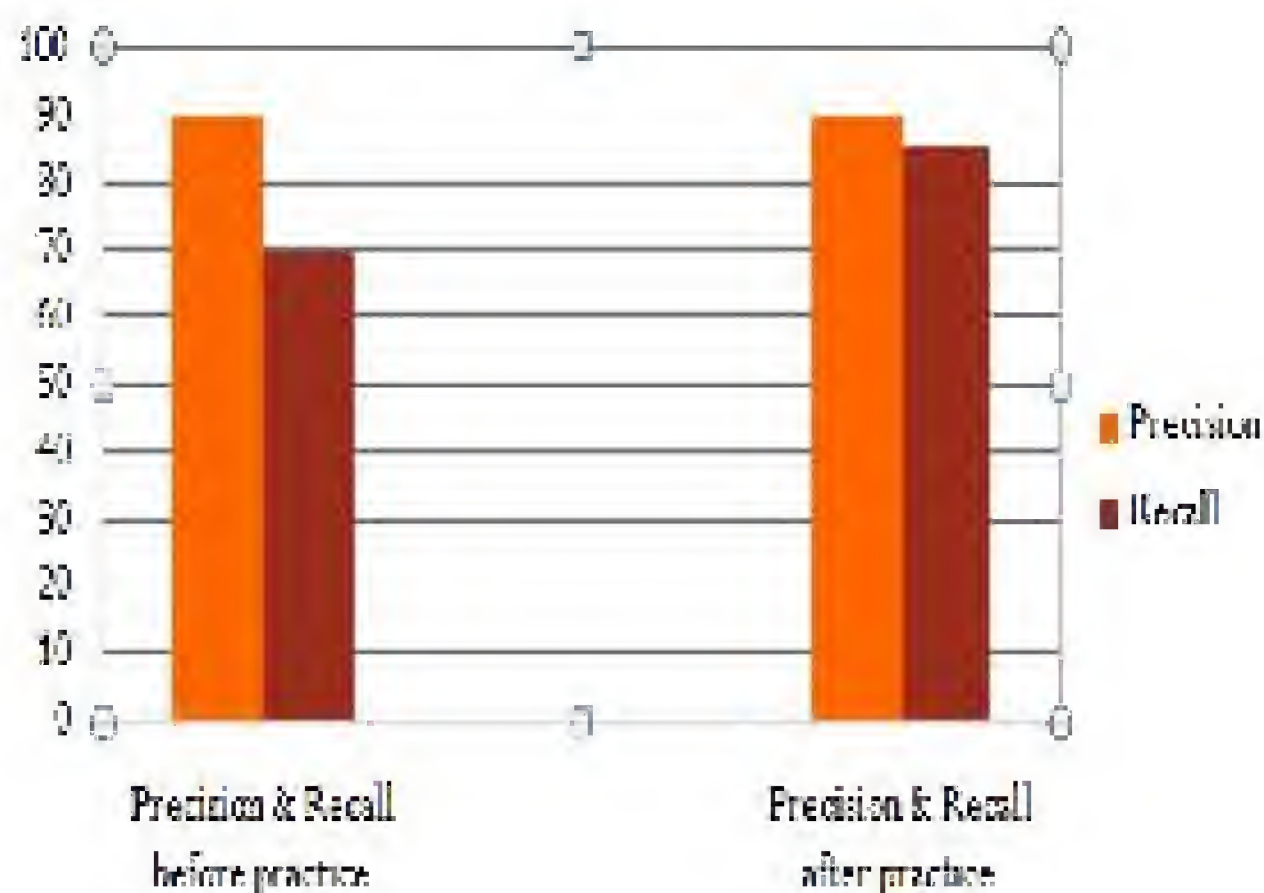


Fig. 4 Precision & Recall chart of VPI

a]Sound Intensity

The frequency of sound is divided in three terms fast, medium and slow. As the result of difference between initial coordinate and last coordinate by difference equation. The result is compare with threshold value. If result value is less than 15 it consider as slower stroke. Similarly, if result is greater that 40 it consider as faster stroke and range between 16 to 39 consider as medium pitch of stroke. The threshold value can be manually changed as per player's expectation.

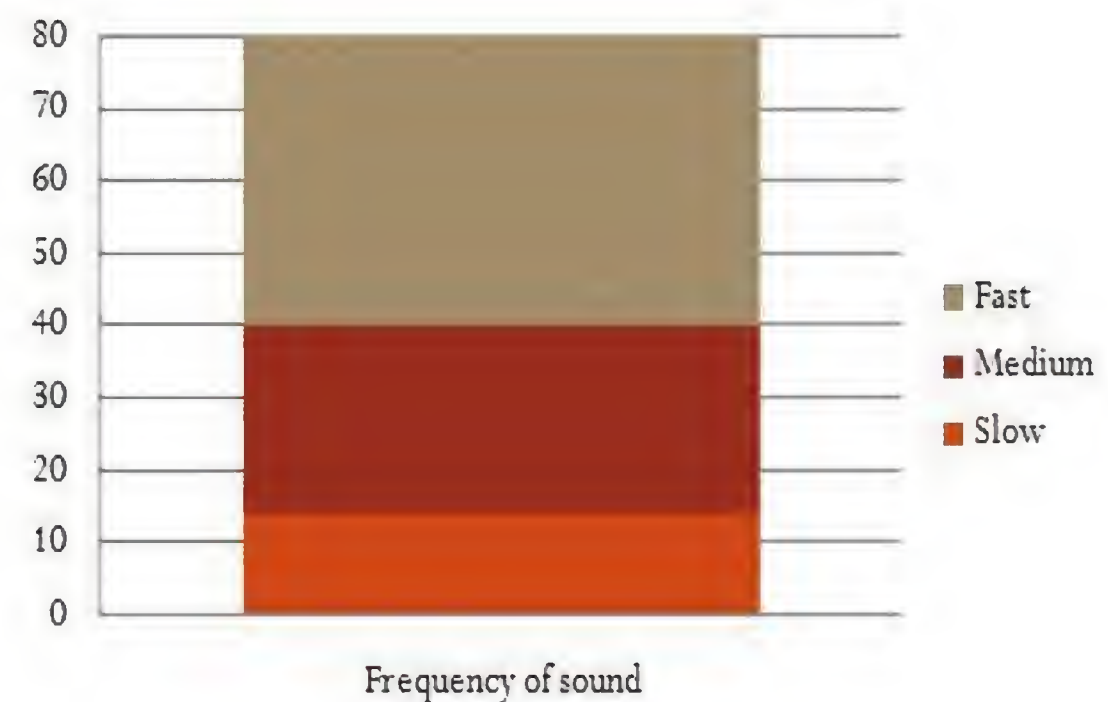


Fig . 5 Sound Intensity of VPI

B. Traditional system VS VPI Performance

The VPI player itself capable of providing customization in terms of sound files.lesser the cost due to only single use of kinect than multiple instrument.for new learner it is necessary to buy all instrument for practice.As VPI player removes need of bulky instrument storage space can be reduce rapidly.we use xbox kinect which can be used for all kinect application as well as VPI player whereas percussion instrument is only used as instruments player.The chart is drawn by information obtained from internet sources.

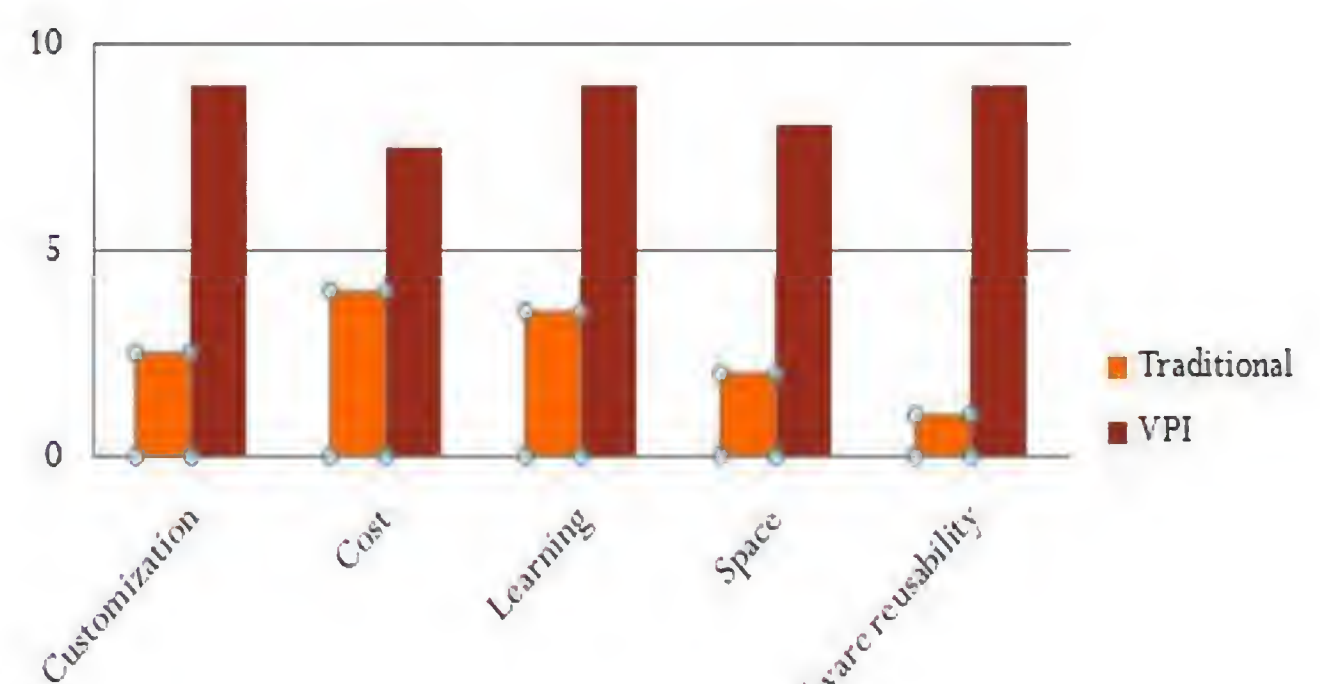


Fig.6 Traditional system VS VPI

V. Conclusion and Future Scope

The paper mainly emphasises on facility provided by Microsoft kinect sensor for hand gesture recognition and classification of percussion instrument. The VPI player proves itself as Human computer reciprocal tool in music world .VPI player fulfil the need of instrument player by eliminating physical instruments, transportation cost of orchestra as well as heavy prices of

instrument. It gives the customization facility for sound as well as reconfiguration of instrument as per player's requirement. For selection of instrument in overlapping window the cosine algorithm is used which map the difference of left and right hand coordinates with instrument coordinates and played the instrument as per the result of mathematical term. The faster and lower stroke on instrument is control by calculating the difference between first stroke and last stroke and comparing the result with threshold value .This threshold value can be change manually which helpful to control the pitch of sound .In future our aim to have system to work on all kind of instrument as well as introducing virtual music teacher concept and compete with the master players in percussion instruments.

ACKNOWLEDGMENT

I take this opportunity to express my deep sense of honour towards my guide Prof. Shabda Dongaonkar for giving me this convenience to select and present this dissertation topic and for encouragement providing me with the best facilities for presentation work. She takes responsibility to continue my dissertation work by giving her quality time to deliver the knowledge .I thankful for her indispensable support, priceless suggestions and for most important time lent as and when required.

REFERENCES

- i. Prateem Chakraborty, Prashant Sarawgi, Ankit Mehrotra, Gaurav Agrawal, Ratika Pradhan, "Hand Gesture Recognition: A comparative Study", in *Proceedings of the International MultiConference of Engineers and Computer scientists* , Vol.1,2008
- ii. Ing-Jr Ding ,Che-Wei Chang and Chang-Jyun He, "A Kinect Based Gesture Command control Method for Human Action Imitations of Humanoid Robots", in *proceedings of 2014 International Conference on Fuzzy Theory and its Applications*,2014.
- iii. YU Bo ,CHEN YongQiang, HUANG YingShu,XIA Chenjie, "Static hand gesture recognition algorithm based on finger angle characteristics", in *Proceedings of the 33rd Chinese Control Conference*, July 2014.
- iv. Ankita Saxena,Deepak Kumar Jain,Ananaya Singhal, "Hand Gesture Recognition using An android device", in *fourth international Conference on Communication System and Network Technologies*,2014.
- v. Chen-Chiung Hsieh and dung-Hua Liou,David Lee, " A Real Time Hand Gesture Recognition System Using Motion History Image" , in *proceedings of the 2nd international Conference on signal Processing System*,2010.

- vi. H.Haggag, M.Hossny, S.Haggag, S.Nahvandi and D.Creighton, "Safety Application Using Kinect Technology",in *2014 IEEE International conference on System, Man,and Cybernetics*, Oct 2014.
- vii. Rama B.Dan, P.S.Method, "Survey on Hand Gesture Recognition approaches", *International Journal of computer science and information technologies*, vol.3, No.4, July 2012.
- viii. Rafiqul Zaman Khan, Noor Adnan Ibraheem, " Hand Gesture Recognition: Literature Review", *International Journal of Artificial Intelligence & application*, Vol.3, No.4, July 2012
- ix. Sunny Amatya, Somrak Petchartee, "Real Time Kinect Based Robotic Arm Manipulation with Five Degree of Freedom",in *Asian Conference on Defense technology*, 2015.
- x. Bogdan Ionescu,Didier Coquin,Patrik Lambert,Vasile Buzuloiu, "Dynamic Hand Gesture Recognition Using the skeleton of the hand", *EURASIP Journal on Applied Signal Processing* 2005.
- xi. Zhou Ren,Junsong Yuan,Jingjing Meng,Zhengyou Zhang, "Robust Part-Based Hand Gesture Recognition Using kinect Sensor", *IEEE TRANSACTION ON MULTIMEDIA*, Vol.15,No.5,2013.
- xii. <https://en.wikipedia.org/wiki/Kinect>, percussion instrument, cosine similarity.



Priti Sanap received her Bachelor's degree in Computer Engineering from SND COE and RC,Pune University,India .Now she is pursuing her M.E degree in Computer Engg. from GHRCEM College, PuneUniversity, Pune, India.



Prof.Shabda Dongaonkar was born in Nanded Maharastra. She received ME degree from pune university .Her research interests are in wireless sensor network and computer network. She is assistance Professor at GHRCEM wagholi for 8.5 years .She published eight journal papers.